

(19) 世界知的所有権機関  
国際事務局



(43) 国際公開日  
2004 年 2 月 19 日 (19.02.2004)

PCT

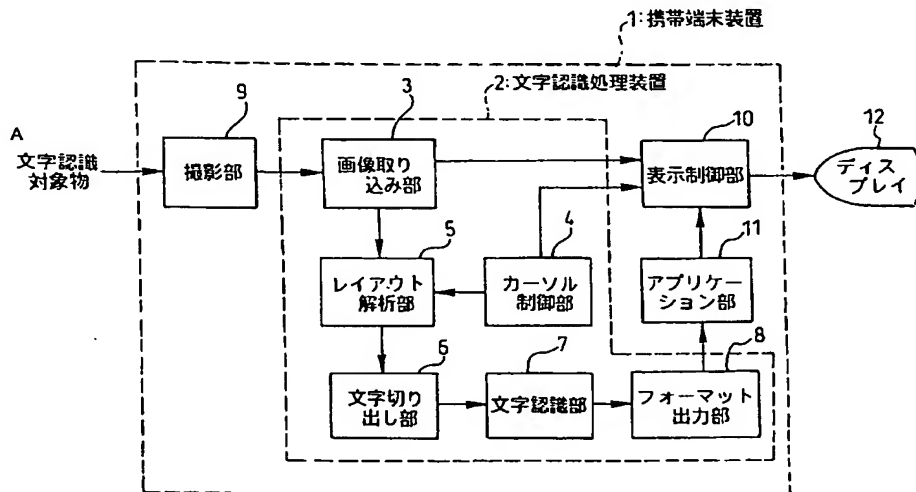
(10) 国際公開番号  
WO 2004/015619 A1

- (51) 国際特許分類: G06K 9/34, 9/03, 9/20, H04M 1/00, 1/56
- (21) 国際出願番号: PCT/JP2003/009483
- (22) 国際出願日: 2003 年 7 月 25 日 (25.07.2003)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願2002-230312 2002 年 8 月 7 日 (07.08.2002) JP  
特願2003-133171 2003 年 5 月 12 日 (12.05.2003) JP
- (71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府 門真市 大字門真 1 0 0 6 番地 Osaka (JP).
- (72) 発明者; および  
(75) 発明者/出願人 (米国についてのみ): 藤田 幸宏 (FUJITA, Yukihiro) [JP/JP]; 〒228-0802 神奈川県 相模原市 上鶴間 5 7 6-1-A 2 0 2 Kanagawa (JP). 沢村 恒治 (SAWAMURA, Kouji) [JP/JP]; 〒241-0002 神奈川県 横浜市 旭区 上白根 3-2 9-1 3-2 0 1 Kanagawa (JP). 伊藤 功将 (ITO, Katsuyuki) [JP/JP]; 〒224-0054 神奈川県 横浜市 都筑区 佐江戸町 7 4 9-1 Kanagawa (JP).
- (74) 代理人: 小栗 昌平, 外 (OGURI, Shohel et al.); 〒107-6028 東京都 港区 赤坂一丁目 1 2 番 3 2 号 アーク森ビル 2 8 階 栄光特許事務所 Tokyo (JP).
- (81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG,

[続葉有]

(54) Title: CHARACTER RECOGNITION PROCESSING DEVICE, CHARACTER RECOGNITION PROCESSING METHOD, AND MOBILE TERMINAL DEVICE

(54) 発明の名称: 文字認識処理装置及び文字認識処理方法、並びに携帯端末装置



1...MOBILE TERMINAL DEVICE  
A...CHARACTER RECOGNITION OBJECT  
9...IMAGING SECTION  
3...IMAGE ACQUISITION SECTION  
2...CHARACTER RECOGNITION PROCESSING DEVICE  
10...DISPLAY CONTROL SECTION  
12...DISPLAY

5...LAYOUT ANALYSIS SECTION  
4...CURSOR CONTROL SECTION  
11...APPLICATION SECTION  
6...CHARACTER ISOLATION SECTION  
7...CHARACTER RECOGNITION SECTION  
8...FORMAT OUTPUT SECTION

(57) Abstract: It is possible to improve the recognition ratio and recognition speed, to simplify the specification required for the imaging device and the devices constituting a system, and to reduce the power consumption and manufacturing cost. A character recognition processing method displays a character image of a character recognition object together with a cursor of a character frame for character recognition on

[続葉有]



SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VC, VN, YU, ZA, ZM, ZW.

OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

(84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ,  
SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM,  
AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許  
(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,  
GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR),

添付公開書類:

— 国際調査報告書

2文字コード及び他の略語については、定期発行される  
各PCTガゼットの巻頭に掲載されている「コードと略語  
のガイダンスノート」を参照。

a display (12) and the character image is imaged by an imaging section (9) with the cursor matched with the character string to be recognized. Image data of the character image imaged is acquired by an image acquisition section (3). A layout analysis section (5) receives the image data of the character image and cursor position information from a cursor control section (4) corresponding to the character image and compares the image data to the cursor position information so as to analyze the line of the character string to be recognized and character arrangement. According to the analysis result, a character isolation section (6) isolates a character. A character recognition section (7) performs character recognition for the character image and converts the image data into character data.

(57) 要約: 本発明の課題は、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減可能とし、低消費電力化及びコストダウンを図ることである。本発明の文字認識処理方法は、文字認識対象物の文字画像を文字認識用文字枠のカーソルとともにディスプレイ(12)に表示し、認識対象文字列にカーソルを合わせた状態で撮影部(9)により撮影する。画像取り込み部(3)において、撮影された文字画像の画像データを取り込み、レイアウト解析部(5)において、取り込んだ文字画像の画像データとこの文字画像に対応するカーソル制御部(4)からのカーソル位置情報とを受け取り、画像データとカーソル位置情報とを照合して認識対象文字列の行や文字の配置を解析する。そして、文字切り出し部(6)で解析結果を基に文字の切り出しを行い、文字認識部(7)によって文字画像に対する文字認識を行って画像データを文字データに変換する。

## 明 細 書

## 文字認識処理装置及び文字認識処理方法、並びに携帯端末装置

## 5 &lt;技術分野&gt;

本発明は、携帯電話機や携帯情報端末等の携帯端末装置に用いて好適な文字認識処理装置及び文字認識処理方法、並びに携帯端末装置に関する。

## &lt;背景技術&gt;

- 10 従来の文字認識方法及び装置として、例えば特開 2 0 0 1 - 6 0 2 5 0 号公報（特許文献 1）や特開 2 0 0 0 - 7 6 3 7 8 号公報（特許文献 2）に開示されているものが知られている。

特開 2 0 0 1 - 6 0 2 5 0 号公報に開示された文字認識方法及び装置は、画像入力された文字から特徴量を抽出し、抽出した特徴量に基づいて文字の認識を行うものである。この従来例では、画像入力された文字を所定間隔で走査する毎に走査線と文字との交差回数を取得するとともに、取得した交差回数の隣接する走査線毎の差分の総和を特徴量の 1 つとして文字の認識を行うものである。認識対象となる文字に対する走査は、少なくとも 2 つの異なる方向に所定間隔に行ったり、直交する 2 方向に所定間隔で行うようにしている。

- 20 他方、特開 2 0 0 0 - 7 6 3 7 8 号公報に開示されている文字認識方法は、入力された画像データの輪郭線データを作成し、作成した輪郭線データを正規化して得たデータを、予め登録している標準文字の輪郭線データを抽出して作成した辞書データに照合して文字認識を行うものである。この際、入力された画像データの書体や歪などによっては、辞書データとの照合前に画像データの輪郭線データを辞書データで認識できる書体に変形したり、あるいは歪が是正されるように変形するようにしている。

[特許文献 1]

特開 2 0 0 1 - 6 0 2 5 0 号公報

[特許文献 2]

- 30 特開 2 0 0 0 - 7 6 3 7 8 号公報

しかしながら、従来の文字認識方法及び装置においては、取り込んだ文字画像データから文字認識処理を行う際に、認識率や認識速度の向上を図るためには、レイアウト解析や行、文字の切り出し作業においてそれぞれ走査線や輪郭線データ等を用いて特別な処理を行う必要があった。

- 5      また、一般に文字認識を行う文字列全体を一度に撮影する場合、1文字に対する解像度を確保するには撮影装置（例えばCCDカメラなど）の画素数を大きくしなければならない上、必要以上に大規模なデータを扱うために、処理に必要なホストCPUやDSPのMIPS (Million Instructions Per Second) 値やワークRAMを大きくしなければならないという問題点もある。
- 10      本発明は、上記事情に鑑みてなされたもので、その目的は、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認識処理方法、並びに携帯端末装置を提供することにある。

#### 15      <発明の開示>

- 本発明に係る文字認識処理装置は、文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手段と、前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手段と、を備えたものである。
- 20

- 上記構成では、文字認識対象物の文字認識を行う際に、認識対象となる文字画像を文字認識用文字枠とともに液晶表示素子等による表示部に表示して、認識対象文字列に文字認識用文字枠のカーソルを合わせた状態で文字画像をCCDカメラ等の撮像部により撮影する。そして、画像取り込み手段で文字画像の画像データを取り込み、レイアウト解析手段により取り込んだ文字画像の画像データに対してカーソル位置情報を照合して行や文字等の配置を解析し、文字切り出し手段
- 25

及び文字認識手段によって文字画像に対する文字の切り出し及び文字認識を行う。

このため、レイアウト解析手段においてカーソル位置情報を利用して認識対象文字列の行や文字の配置情報を正確に解析することが可能となる。これにより、認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することで文

- 5 字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能であるため、撮影部を多画素化するなどの高性能の仕様は必要なく、装置に要求されるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減することができ、これによって、CPUやDSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できるため、装置各部の仕様の程度を低減でき、
- 10 低消費電力化及びコストダウンを図れる。

また、上記構成において、前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを取り込み、前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごとに、前記カーソル位置情報を照

- 15 合して文字の配置を解析するものとする。

上記構成では、文字認識対象物の認識対象文字列が大きい場合など、1度の撮影では文字画像全体を撮影できない場合に、撮影部を移動させながら連続撮影を行って文字画像を撮影する。そして、得られた文字画像の画像データについて、画像取り込み手段で文字画像全体から所定領域ずつ文字画像の画像データを取り

20 込み、レイアウト解析手段により取り込んだ複数の文字画像の画像データごとにカーソル位置情報を照合して行や文字等の配置を解析する。このように文字画像をいくつかのコマの静止画像に分けてカーソル位置情報とともに取り込むことにより、認識対象文字列が長い場合などでもレイアウト解析手段においてカーソル位置情報を利用して認識対象文字列の配置情報を正確に解析することが可能となる。

25

また、上記構成において、前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを取り込み、前記レイアウト解析手段は、前記取り込んだ複数の文字画像を連結した画像データについて、前記カーソル位

置情報を照合して文字の配置を解析するものとする。

上記構成では、文字認識対象物の認識対象文字列が大きい場合など、1度の撮影では文字画像全体を撮影できない場合に、撮影部を移動させながら連続撮影を行って文字画像を撮影する。そして、得られた文字画像の画像データについて、

- 5 画像取り込み手段で文字画像全体から所定領域ずつ文字画像の画像データを取り込み、レイアウト解析手段により取り込んだ複数の文字画像の画像データを一旦1つの画像に連結した後、カーソル位置情報を照合して行や文字等の配置を解析する。このように文字画像をいくつかのコマの静止画像に分けてカーソル位置情報とともに取り込むことにより、認識対象文字列が長い場合などでもレイアウト
- 10 解析手段においてカーソル位置情報を利用して認識対象文字列の配置情報を正確に解析することが可能となる。

また本発明は、上記いずれかに記載の文字認識処理装置を備えた携帯端末装置を提供する。

- 上記構成によれば、文字認識処理装置を搭載した携帯端末装置において、文字
- 15 認識の認識率及び認識速度の向上を図りつつ、撮影部やCPU、DSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できるため、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

- 本発明に係る携帯端末装置は、文字認識用の認識対象となる文字画像を撮影する撮影手段と、前記撮影された文字画像の画像データを取り込む画像取り込み手段と、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手段と、前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手段と、を備えたものである。
- 20
- 25

上記構成によれば、文字画像の画像データと文字認識用文字枠の位置を示すカーソル位置情報とに基づき、文字認識用文字枠に対応する文字のレイアウト解析を行って文字認識処理を実行することによって、文字認識の認識率及び認識速度の向上を図りつつ、撮影部やCPU、DSP、メモリ等のシステムを構成するデ

バイスに必要とされるスペックを軽減できる。このため、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

また、上記構成において、前記撮影手段による撮影時に、前記文字画像と重ねて文字認識用文字枠を表示する文字枠表示手段を備えるものとする。

- 5      この構成によれば、使用者は表示された文字認識用文字枠を見ながら容易に認識対象の文字を文字認識用文字枠に合わせて撮影し、文字認識を行うことが可能となる。

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字情報を表示する認識文字表示手段を備えるものとする。

- 10     この構成によれば、使用者は認識結果の文字情報を表示によって容易に確認可能となる。

また、上記構成において、前記認識文字表示手段は、前記認識結果の文字情報を所定の文字単位ごとに個別に選択可能に表示するものとする。

- 15     この構成によれば、使用者は表示された文字情報を基に、認識した文字情報を文字種別などの単位ごとに個別に選択して活用することが可能となる。

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字情報を保存する認識文字保存手段を備えるものとする。

この構成によれば、使用者は認識した文字情報を所望の形態で保存して活用することが可能となる。

- 20     また、上記構成において、前記認識文字保存手段は、前記文字情報を認識文字格納領域に保存するものであることとする。

この構成によれば、認識文字格納領域において例えば複数の認識結果の文字情報を格納して利用することが可能となる。

- 25     また、上記構成において、前記認識文字保存手段は、前記文字情報の種別が電話番号、メールアドレス、URL (Uniform Resource Locator) のいずれかであった場合に、その文字情報の種別に応じたデータベースに登録するものであることとする。

この構成によれば、電話番号、メールアドレス、URLなどを文字認識した場合に、その文字情報の種別に応じて電話帳やブックマークなどのデータベースへ

登録して利用することが可能となる。

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字情報を当該文字情報の種別に応じて活用する認識文字活用手段を備えることとする。

- 5      この構成によれば、認識した文字情報の種別に応じて、例えば電話番号、メールアドレス、URLなどの種別毎に文字情報を使用者が有効に利用することが可能となる。

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別が電話番号であった場合にその電話番号への発信画面を表示するものとする。

- 10     この構成によれば、認識した文字情報が電話番号であった場合にその電話番号への発信画面を表示し、使用者は簡便に発信動作を行うことが可能となる。

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別がメールアドレスであった場合にそのメールアドレスへの電子メール作成画面を表示するものとする。

- 15     この構成によれば、認識した文字情報がメールアドレスであった場合にそのメールアドレスへの電子メール作成画面を表示し、使用者は簡便に電子メールの作成、送信動作を行うことが可能となる。

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別がURL (Uniform Resource Locator) であった場合にそのURLへのネットワーク

- 20     接続画面を表示するものとする。

この構成によれば、認識した文字情報がURLであった場合にそのURLへのネットワーク接続画面を表示し、使用者は簡便にインターネット等のネットワークへの接続動作を行って目的の情報を得ることが可能となる。

- 25     また、上記いずれかの構成において、前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数取得した場合に、これらの複数の文字画像の画像データを連結して取り込み、前記文字認識手段は、前記連結された文字画像の画像データの文字認識を行って文字情報に変換するものとする。

この構成によれば、認識対象の文字列が1回の撮影範囲に収まらない場合に、



連続的に文字画像を文字認識用文字枠と対応づけて撮影し、得られた複数の文字画像を連結することで、長い文字列であっても文字認識の認識率及び認識速度の向上を図ることが可能である。

5 また、上記いずれかの構成において、前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数取得した場合に、これらの複数の文字画像の画像データを取り込み、前記文字認識手段は、前記取り込まれた複数の文字画像のそれぞれについて画像データの文字認識を行って文字情報に変換するものとする。

10 この構成によれば、認識対象の文字列が1回の撮影範囲に収まらない場合に、連続的に文字画像を文字認識用文字枠と対応づけて撮影し、得られた複数の文字画像のそれぞれについて画像データの文字認識を行うことで、長い文字列であっても文字認識の認識率及び認識速度の向上を図ることが可能である。

15 また、上記いずれかの構成において、前記文字認識手段は、認識対象の文字種別に応じた複数の認識モードを有し、予め設定された認識モードによって対応する文字種別に適した文字認識処理を行うものとする。

この構成によれば、使用者が文字認識しようとする文字種別に対応した認識モードを設定することで、その文字種別に適した文字認識処理を行うことができ、認識率を高めることが可能となる。

20 本発明に係る文字認識処理方法は、認識対象となる文字画像を文字認識用文字枠とともに表示して撮影するステップと、前記撮影した文字画像の画像データを取り込むステップと、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するステップと、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するステップと、前記文字配置の解析結果を基に文字画像を切り出すステップと、前記切り出された文字画像を文字として認識し文字情報に変換するステップと、を有することを特徴とする。

この方法によれば、カーソル位置情報を利用して認識対象文字列の行や文字の配置情報を正確に解析することが可能となり、認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することにより、装置各部に要求されるスペ

ックの軽減が図れるとともに、装置内でのデータ処理量を低減できるため、装置の仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能となる。

また本発明は、上記の各ステップをコンピュータに実行させるための文字認識  
5 処理プログラムを提供する。

このプログラムによれば、文字画像の画像データと文字認識用文字枠の位置を示すカーソル位置情報とに基づき、文字認識用文字枠に対応する文字のレイアウト解析を正確に行って文字認識処理を実行でき、文字認識の認識率及び認識速度の向上を図りつつ、装置各部の仕様の程度を低減でき、低消費電力化及びコスト  
10 ダウンを図れる。

#### <図面の簡単な説明>

図 1 は、本発明の第 1 実施形態に係る文字認識処理装置を備えた携帯端末装置の構成を示すブロック図であり、

15 図 2 は、本発明の第 1 実施形態に係る文字認識処理装置における文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示す図であり、

図 3 は、本発明の第 2 実施形態に係る文字認識処理装置における文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示す図であり、

図 4 は、本発明の第 2 実施形態に係る文字認識処理装置における文字認識処理  
20 時の動作説明図であり、複数に分割した認識対象文字列に対する文字認識処理の様子を示す図であり、

図 5 は、本発明の第 3 実施形態に係る携帯電話装置の外観構成を示す図であり、  
(A) は正面図、(B) は背面図であり、

図 6 は、第 3 実施形態に係る携帯電話装置の構成を示すブロック図であり、

25 図 7 は、第 3 実施形態の携帯電話装置における文字認識処理に関する機能的構成を示すブロック図であり、

図 8 は、図 7 の構成における文字利用部の機能的構成を示す説明図であり、

図 9 は、第 3 実施形態における文字認識及び文字利用に関する全体の処理の流れを示す説明図であり、

図 1 0 は、文字認識対象物の一例を示す図であり、

図 1 1 は、第 3 実施形態における OCR 機能のメインメニュー画面の一例を示す図であり、

図 1 2 は、第 3 実施形態における OCR 機能の各画面の一例を示したもので、

5 (A) はプレビュー表示画面、(B) はカメラ機能設定画面、(C) は認識モード設定画面をそれぞれ示す図であり、

図 1 3 は、第 3 実施形態における OCR 機能の各画面の一例を示したもので、

(A) は認識文字表示画面、(B) は機能メニュー表示画面、(C) は音声発信画面、(D) はインターネット接続画面、(E) はメール作成画面をそれぞれ示

10 す図である。

なお、図中の符号、1 は携帯端末装置、2 は文字認識処理装置、3 は画像取り込み部、4 はカーソル制御部、5 はレイアウト解析部、6 は文字切り出し部、7 は文字認識部、8 はフォーマット出力部、9 は撮影部、1 0 は表示制御部、1 1 はアプリケーション部、1 2 はディスプレイ、3 0 は携帯電話装置、3 3 は第 1  
15 表示部、3 4 はカメラ、3 5 は第 2 表示部、3 6 は操作キー、4 1 は撮影部、4 2 は操作部、4 3 は表示部、4 4 は通信部、4 5 は音声処理部、4 6 は記憶部、4 7 は制御処理部、5 1 は文字認識部、5 2 は文字利用部、6 1 は仮保存部、6 2 は保存部、6 3 は出力部、6 4 は種別判定部である。

## 20 <発明を実施するための最良の形態>

以下、図面を参照して本発明の実施形態を説明する。

### (第 1 実施形態)

図 1 は、本発明の第 1 実施形態に係る文字認識処理装置を備えた携帯端末装置の構成を示すブロック図である。本実施形態では、文字認識処理装置 2 を携帯端  
25 末装置 1 に搭載した構成例を説明する。

本実施形態の文字認識処理装置 2 は、画像取り込み部 3 と、カーソル制御部 4 と、レイアウト解析部 5 と、文字切り出し部 6 と、文字認識部 7 と、フォーマット出力部 8 とを備えて構成される。これらの構成要素は、プロセッサ及びメモリ等を有してなる制御処理部におけるソフトウェアプログラムの動作、またはソフ

トウェアとハードウェアの組み合わせ、もしくは専用のハードウェアによって実現される。

携帯端末装置 1 は、文字認識処理装置 2 とともに、例えば CCD カメラ等からなる撮影部 9 と、表示制御部 10 と、アプリケーション部 11 とを有している。

- 5 撮影部 9 は、文字認識対象物を撮影して撮影データとして取り込み、文字認識処理装置 2 の画像取り込み部 3 に入力する。画像取り込み部 3 は、画像取り込み手段に相当するもので、入力された撮影データを所定のフォーマットの画像データに変換して表示制御部 10 に入力する。カーソル制御部 4 は、カーソル情報出力手段に相当するもので、文字認識用文字枠として機能するカーソルに関するカーソル位置情報を表示制御部 10 に入力する。表示制御部 10 は、入力された画像データとカーソル位置情報とを表示データに変換し、さらに液晶表示素子等からなるディスプレイ 12 に適したフォーマットに加工して出力し、ディスプレイ 12 において情報表示を行う。

- レイアウト解析部 5 は、レイアウト解析手段に相当するもので、文字認識時に  
15 画像取り込み部 3 から所定フォーマットの画像データを受け取るとともに、カーソル制御部 4 からカーソル位置情報を受け取り、カーソル位置情報及び画像データを照合してレイアウト解析を行う。文字切り出し部 6 は、文字切り出し手段に相当するもので、レイアウト解析結果を元に画像データを文字単位で切り出して抽出する。文字認識部 7 は、文字認識手段に相当するもので、切り出された文字  
20 単位の画像データの認識処理を行って文字データに変換する。フォーマット出力部 8 は、文字認識部 7 で認識された文字データを適切なフォーマットに変換する。アプリケーション部 11 は、フォーマット出力部 8 でフォーマット変換された文字データをアプリケーション用のデータとして利用し、必要に応じて表示制御部 10 に入力してディスプレイ 12 に表示する。

- 25 次に、図 2 を参照して第 1 実施形態の文字認識処理装置の動作について説明する。図 2 は第 1 実施形態に係る文字認識処理時の動作説明図であり、認識対象文字列を撮影している状態を示している。

携帯端末装置 1 の撮影部 9 によって文字認識対象物 16 の中の認識対象文字列 17 を撮影し、文字認識を行う。撮影している認識対象文字列 17 をディスプレ

イ 1 2 に表示する際に、認識対象文字列 1 7 と併せて文字認識用文字枠であるカーソル 1 8 が表示される。使用者は撮影位置を調整しながら認識対象文字列 1 7 をカーソル 1 8 に合わせ、この状態で操作ボタン 1 9 を操作して静止画像の撮影を実行する。

- 5      撮影された認識対象文字列 1 7 の静止画像は、画像取り込み部 3 で定められたフォーマットの画像データに変換されて取り込まれ、レイアウト解析部 5 に渡される。また、レイアウト解析部 5 ではカーソル制御部 4 からカーソル位置情報を受け取り、カーソル位置情報と文字画像の画像データとを照合する。このように、レイアウト解析部 5 では、カーソル位置情報を利用して認識対象文字列 1 7 の行  
10   及び文字の配置情報を正確に解析することができる。そして、文字切り出し部 6 で画像データの文字単位での切り出しが行われ、文字認識部 7 において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。

- このように、第 1 実施形態の文字認識処理装置では、文字認識対象物 1 6 の認識対象文字列 1 7 を文字認識する際に、ディスプレイ 1 2 に認識対象文字列 1 7  
15   の画像と併せて文字認識用文字枠のカーソル 1 8 を表示させておき、使用者が認識対象文字列 1 7 にカーソル 1 8 を合わせた状態で静止画像の撮影を行い、撮影された文字画像の画像データをカーソル位置情報とともに取り込むようにする。これにより、レイアウト解析部 5 においてカーソル位置情報を利用して行や文字の配置情報を容易かつ正確に解析することができ、文字認識における認識率及び  
20   認識速度の向上を図ることができる。

### (第 2 実施形態)

- 次に、図 3 及び図 4 を用いて本発明の第 2 実施形態に係る文字認識処理装置の動作について説明する。図 3 及び図 4 は第 2 実施形態に係る文字認識処理時の動作説明図であり、図 3 は認識対象文字列を撮影している状態を示し、図 4 は複数  
25   に分割した認識対象文字列に対する文字認識処理の様子を示している。なお、文字認識処理装置の構成は第 1 実施形態と同様であるので、図 1 を援用するものとする。

第 2 実施形態では、第 1 実施形態と同様、携帯端末装置 1 の撮影部 9 によって

文字認識対象物 2 2 の中の認識対象文字列 2 3 を撮影して文字認識を行う。この場合、認識対象文字列 2 3 の範囲が広いために 1 度に取り込めない場合は以下のよう

5 撮影している認識対象文字列 2 3 をディスプレイ 1 2 に表示する際に、認識対象文字列 2 3 と併せて文字認識用文字枠であるカーソル 1 8 が表示される。使用者は撮影位置を調整しながら認識対象文字列 2 3 の開始端位置の部分（図 3 の例では左端の部分）をカーソル 1 8 に合わせ、この状態で操作ボタン 1 9 を操作して連続静止画像撮影を開始する。撮影開始後、使用者は携帯端末装置 1 の移動方向を固定しながら、文字の筆記方向に向かうように撮影開始時の開始端位置と逆

10 方向（図 3 の例では右方向）に認識対象文字列 2 3 に対して平行にスライドさせ、認識対象文字列 2 3 の終了端位置（図 3 の例では右端位置）までカーソル 1 8 に収まるようにディスプレイ 1 2 に表示させたところで連続静止画像撮影を終了する。このとき、使用者は例えば連続静止画像撮影の最中は操作ボタン 1 9 を押下したままにする。

15 撮影された認識対象文字列 2 3 の連続静止画像は、画像取り込み部 3 で図 4 のように複数の画像データに分割されて取り込まれる。この場合、文字画像全体から 1 行分の文字画像についていくつかのコマ（1）～（4）の静止画像に分割されて抽出された後、定められたフォーマットの画像データに変換されてレイアウト解析部 5 に渡される。

20 また、レイアウト解析部 5 ではカーソル制御部 4 からカーソル位置情報を受け取り、カーソル位置情報と文字画像の画像データとを照合する。但し、レイアウト解析に利用する各コマの文字画像に関してそれぞれの文字がカーソルの文字枠に収まっていない可能性があるため、カーソル位置情報としては「行」に対する配置情報の解析には問題なく利用することができるが、文字ごとの配置情報の解析

25 には利用できない場合がある。この場合、カーソル位置情報を連続静止画像の移動方向にシフトさせることによって各文字の配置情報を解析することも可能である。このように、レイアウト解析部 5 においては、認識対象文字列 2 3 が長い場合でもカーソル位置情報を利用して認識対象文字列 2 3 の配置情報を正確に解析することができる。本実施形態では、行の配置情報について正確な解析が可能

である。

そして、文字切り出し部 6 で文字単位の切り出しが行われ、文字認識部 7 において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。この場合、文字認識結果として、コマ (1) では「A B C D E F G H I J K L」が得られ、コマ (2) では「F G H I J K L M N O P」が得られ、コマ (3) では「J K L M N O P Q R S T U」が得られ、コマ (4) では「O P Q R S T U V W」が得られる。文字認識部 7 では、これらのコマ毎の文字認識結果を合成することにより、最終文字認識結果として「A B C D E F G H I J K L M N O P Q R S T U V W」が得られる。

- 10     なお、第 2 実施形態では、簡単のため 1 行分の文字画像をカーソルと照合して一連の文字認識の処理をするかのごとく説明したが、図 3 に示すようにカーソル 1 8 が 2 行分存在する場合、一連の文字認識の処理についても 2 行分同時に行うことができる。

- 15     このように、第 2 実施形態の文字認識処理装置では、文字認識対象物 2 2 の認識対象文字列 2 3 を文字認識する際に、ディスプレイ 1 2 に認識対象文字列 1 7 の画像と併せて文字認識用文字枠のカーソル 1 8 を表示させておき、使用者が携帯端末装置 1 を移動させながら連続静止画像撮影を行う。そして、撮影した文字画像をいくつかの静止画像に分けて、各コマの画像データをカーソル位置情報とともに取り込むようにする。これにより、レイアウト解析部 5 においてカーソル  
20     位置情報を利用して行の配置情報を容易かつ正確に解析することができ、文字認識における認識率及び認識速度の向上を図ることができる。

- 25     なお、上記の第 2 実施形態においては、連続静止画像撮影を行って複数の画像データを得る際に、携帯端末装置 1 の操作ボタン 1 9 を押下したままで認識対象文字列上を移動させるようにしているが、これに限るものではない。例えば、使用者が携帯端末装置 1 本体を動かして撮影範囲を移動させながらその都度操作ボタン 1 9 を押下することで断続的に複数の画像データを取得し、認識対象文字列の画像として連続した広い範囲の画像データを得ることも可能である。

また、レイアウト解析部 5 は、画像取り込み部 3 によって取り込まれた複数の文字画像（静止画像）について、各文字画像毎にレイアウト解析してもよいし、

複数の文字画像を一旦１つの文字画像に連結した後レイアウト解析してもよい。

上述したように本実施形態では、カーソル位置情報を利用することで文字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能であるため、撮影部 9 を多画素化するなどの高性能の仕様は必要なく、装置に要求されるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減することができ、これによって、CPU や DSP、メモリ 等のシステムを構成するデバイスに必要とされるスペックを軽減できる。したがって、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

### 10 (第 3 実施形態)

次に、第 3 実施形態として、携帯電話装置に文字認識処理装置を搭載した場合の構成及び動作の一例を示す。

図 5 は本発明の第 3 実施形態に係る携帯電話装置の外観構成を示す図であり、(A) は正面図、(B) は背面図である。本実施形態の携帯電話装置 30 は、第 1 筐体 31 と第 2 筐体 32 とを有して構成される。第 1 筐体 31 の正面側、すなわち使用者が第 1 筐体 31 及び第 2 筐体 32 を開いて把持したときに主に使用者側に向く面には、液晶表示素子等からなる第 1 表示部 33 が設けられている。また、第 1 筐体 31 の正面とは反対側の背面側には、撮影部としてのカメラ 34 と、液晶表示素子等からなる第 2 表示部 35 とが設けられている。

第 2 筐体 32 の正面側には、オンフック／オフフックキー、数字キー、機能キーなどを含む複数のキーボタンからなる操作部としての操作キー 36 が配設されている。機能キー PF1 ～ PF5 には、装置の動作モードや表示画面などによって各種の機能が割り当てられるようになっている。

図 6 は第 3 実施形態に係る携帯電話装置の構成を示すブロック図である。携帯電話装置は、文字認識対象物等の画像を撮影する撮影手段の機能を有する撮影部 41、使用者が操作入力を行う操作部 42、文字認識結果等の各種情報を表示する文字枠表示手段及び認識文字表示手段の機能を有する表示部 43、移動体通信システムの基地局等との間で無線通信を行う通信部 44、音声通話や着信音等の音声処理を行う音声処理部 45、各種情報を記憶する認識文字保存手段の機能を



有する記憶部 4 6、文字認識処理動作や認識文字データの利用などのアプリケーション処理や各部の制御などを行う制御処理部 4 7 を有して構成される。

携帯電話装置で音声通話を行う際は、通信部 4 4、音声処理部 4 5 及び制御処理部 4 7 によって通話音声の入力、送受信、出力を行う。電子メール等のメッセージのやり取りを行ったり、インターネット等のネットワークにアクセスして各種情報を取得する際は、操作部 4 2、表示部 4 3、通信部 4 4 及び制御処理部 4 7 によって文字データの入力、文字データや画像データ等の情報の送受信、表示を行う。画像撮影を行う際は、撮影部 4 1 及び制御処理部 4 7 によって被写体を撮影して撮影画像の画像データを取得する。記憶部 4 6 は、通信相手の電話番号やメールアドレス等の電話帳登録情報、電子メール等の文字データ、撮影画像等の画像データ、着信音等の音声データなどの各種情報を記憶する。

図 7 は第 3 実施形態の携帯電話装置における文字認識処理に関する機能的構成を示すブロック図である。第 3 実施形態では、第 1 及び第 2 実施形態と同様の撮影部 4 1 及び文字認識部 5 1 を備えると共に、認識結果の文字データを各部の機能において利用する認識文字活用手段の機能を有する文字利用部 5 2 を備えている。これらの文字認識部 5 1 及び文字利用部 5 2 は、制御処理部 4 7 及び記憶部 4 6 において、プロセッサ及びメモリを含むハードウェアとこれらを用いて動作するソフトウェアプログラムとによって実現される。

図 8 は図 7 の構成における文字利用部の機能的構成を示す説明図である。文字利用部 5 2 は、仮保存部 6 1、保存部 6 2、出力部 6 3 を有する。仮保存部 6 1 は、認識文字保存手段であり、文字認識処理後の文字データを一時記憶するものである。保存部 6 2 は、認識文字保存手段であり、認識文字格納領域として複数の保存領域を有し、複数の文字データをそれぞれ格納して保存するものである。出力部 6 3 は、文字データの種別を判別する種別判別部 6 4 を有し、認識結果の文字データの種別に応じた出力動作を行うものである。文字認識によって得られた文字データが電話番号の場合は、その電話番号への音声発信またはデータベースである電話帳登録に関する動作を行う。また、メールアドレスの場合は、そのメールアドレスへのメール作成または電話帳登録に関する動作を行う。また、URL (Uniform Resource Locator) の場合は、その URL へアクセスするインタ

ーネット接続またはデータベースであるブックマーク登録に関する動作を行う。

次に、第3実施形態における文字認識及び文字利用に関する動作の具体例について説明する。図9は文字認識及び文字利用に関する全体の処理の流れを示す説明図である。この図9は、表示部43に表示する表示画面を中心に各動作状態の遷移を示したものである。ここでは、図10に示すような文字認識対象物である名刺101に記載された文字を認識対象として撮影し、携帯電話装置に搭載したOCR (Optical Character Recognition) 機能によって文字認識を行う場合を説明する。

携帯電話装置のOCR機能を起動すると、初期画面に相当するメインメニュー画面71が表示される。図11はメインメニュー画面71の一例を示したものである。このメインメニュー画面71には、複数の文字データ登録内容が一覧表示される。本実施形態では、認識文字列の文字データを複数（例えば20個）登録可能となっている。ここで、使用者が「未登録」の欄を選択してハイライト表示等によってフォーカスが当たった状態（図において斜線で示す）とし、機能キーPF1に割り当てられた「選択」ボタンを押下すると、新規の認識対象文字列画像の撮影のため、撮影部41が起動してカメラ撮影待機状態となり、プレビュー表示画面72が表示される。なお、フォーカスを当てる位置を移動して選択対象を切り換える場合は、機能キーPF1の外周部に割り当てられた十字キーを操作することで行う。

図12(A)はプレビュー表示画面72の一例を示したものである。このプレビュー表示画面72には、撮影部41で撮影する認識対象の文字画像81が表示され、この文字画像81上に文字認識領域を示す文字認識用文字枠82が重畳表示される。使用者が文字画像81を見ながら認識対象文字列が文字認識用文字枠82内に収まるように撮影範囲を調整し、機能キーPF1に割り当てられた「撮影」ボタンを押下すると、文字画像81の撮影が行われ、文字認識部51によって認識対象文字列のレイアウト解析、文字切り出し、文字認識等の第1実施形態で述べた一連の文字認識処理が行われる。

また、図12(A)のプレビュー表示画面72において使用者が機能キーPF3に割り当てられた「機能」ボタンを押下すると、撮影部41の機能設定モード

に移行し、カメラ機能設定画面 7 3 が表示される。図 1 2 (B) はカメラ機能設定画面 7 3 の一例を示したものである。このカメラ機能設定画面 7 3 では、設定対象として、認識モード、明暗調整、フラッシュ（撮影用照明）のいずれかを選択できるようなっている。ここで、使用者が「認識モード」の欄を選択して  
5 フォーカスが当たった状態とし、機能キー P F 1 に割り当てられた「選択」ボタンを押下すると、認識モード設定画面 7 4 が表示される。

図 1 2 (C) は認識モード設定画面 7 4 の一例を示したものである。この認識モード設定画面 7 4 では、認識モードとして、自動、URL、メールアドレス、電話番号の各モードのいずれかが選択されると、撮影された文字画像の全ての文  
10 字列に対して、選択されたモードに特化した文字認識処理を行う。フリーモードが選択されると、任意の文字列を対象として文字認識を行い、認識可能な文字列が全て認識文字データとして出力される。自動モードが選択されると、上記 4 つのモードの中から文字認識部 5 1 が文字画像 8 1 の内容に応じて自動的に判定して適切な認識モードで文字認識を行う。ここで、使用者が認識モードを選択して  
15 フォーカスが当たった状態とし、機能キー P F 1 に割り当てられた「選択」ボタンを押下すると、プレビュー表示画面 7 2 に戻る。なお、上記動作では、撮影された文字画像の全ての文字列に対して、選択されたモードに特化した文字認識処理を行うようにしたが、自動、URL、メールアドレス、電話番号の各モードのいずれかが選択された場合に、URL などの対応する文字列のみを認識文字データとして出力するようにしても良い。  
20

プレビュー表示画面 7 2 の状態から撮影及び文字認識処理が完了すると、認識文字表示画面 7 5 が表示される。図 1 3 (A) は認識文字表示画面 7 5 の一例を示したものである。この認識文字表示画面 7 5 では、認識結果の文字列の文字データを表示する。ここで、認識文字列に電話番号、メールアドレス、URL が含まれる場合は、それぞれに下線が付されたアンカー表示がなされ、選択文字列に  
25 フォーカスが当たった状態となる。本実施形態では、認識文字列を解析して電話番号、メールアドレス、URL を識別し、それぞれを個別に選択可能であり、文字列の種別に応じた利用ができるようになっている。

電話番号、メールアドレス、URL の識別には、例えば以下の判定方法を用い

る。

(1) 電話番号

- ・ 0 から始まる 10 桁以上の数字が存在する場合

(2) メールアドレス

- 5     ・ 「@」を含んだ文字が存在する場合

(3) URL

- ・ 「http://」または「https://」から始まる文字が存在する場合

10     なお、上記メインメニュー画面 7 1 において、使用者が登録済みの文字列の欄を選択してフォーカスが当たった状態とし、機能キー P F 1 に割り当てられた「選択」ボタンを押下した場合は、認識文字表示画面 7 5 が表示され、過去に認識されて保存された文字列の文字データが表示される。

15     認識文字表示画面 7 5 において使用者が機能キー P F 3 に割り当てられた「機能」ボタンを押下すると、機能メニュー表示画面 7 6 が表示される。図 1 3 (B) は機能メニュー表示画面 7 6 の一例を示したものである。機能メニュー表示画面 7 6 では、連結、編集、コピー、保存、電話帳新規登録、電話帳追加登録、ブックマーク登録のいずれかを選択できるようになっている。

20     一回で撮影できない長い文字列について続けて撮影し、認識結果を一つにつなぎ合わせる場合は、機能メニュー表示画面 7 6 で「連結」の機能を選択する。使用者が「連結」の欄を選択してフォーカスが当たった状態とし、機能キー P F 1 に割り当てられた「選択」ボタンを押下すると、再度プレビュー表示画面 7 2 が表示され、撮影可能な状態となる。そして前回に続く認識対象文字列の撮影及び文字認識処理が完了すると、認識文字表示画面 7 5 に切り替わり、今回認識された文字列が追加して表示される。なお、文字認識については、複数の画像毎に行ってもよいし、複数の画像を 1 つの画像に連結してから行ってもよい。

25     機能メニュー表示画面 7 6 において使用者が「連結」以外の各機能を選択して機能キー P F 1 に割り当てられた「選択」ボタンを押下すると、それぞれの機能に対応する機能画面 7 7 が表示され、各機能が実行可能となる。「編集」を選択した場合は、テキストエディタが起動して編集画面が表示される。使用者は編集画面において誤認識された文字について修正を行うことができる。「コピー」を

選択した場合は、コピー機能が呼び出されてコピー範囲選択画面が表示され、使用者は範囲指定した文字データをコピーして他の文字入力画面などで利用することができる。「保存」を選択した場合は、認識文字列の文字データの保存処理が行われ、完了すると「保存しました」などの保存完了通知画面が表示される。これにより、メインメニュー画面 7 1 で選択した「未登録」の欄に今回の文字認識処理により得られた認識文字列が格納される。

また、「電話帳新規登録」または「電話帳追加登録」を選択した場合は、電話帳登録画面が表示され、使用者は電話番号やメールアドレス、名前などの認識文字列を電話帳の所望の欄に登録することができる。「ブックマーク登録」を選択した場合は、ブックマーク登録画面が表示され、使用者はURLの認識文字列をブックマークとして登録することができる。

認識文字表示画面 7 5 において、使用者が電話番号、メールアドレス、URL のいずれかの文字列を選択した状態で機能キー P F 1 に割り当てられた「選択」ボタンを押下すると、それぞれの文字列の種別に応じた利用画面 7 8 が表示される。電話番号を選択した場合は、図 1 3 (C) に示すように利用画面として音声発信画面 7 8 a が表示される。この音声発信画面 7 8 a では、文字認識した電話番号宛の音声発信を簡易な操作で行うことが可能である。URL を選択した場合は、図 1 3 (D) に示すように利用画面としてインターネット接続画面 7 8 b が表示され、文字認識したURL によるインターネット接続を簡易な操作で行うことが可能である。メールアドレスを選択した場合は、図 1 3 (E) に示すように利用画面としてメール作成画面 7 8 c が表示され、文字認識したメールアドレス宛の電子メール作成を簡易な操作で行うことが可能である。

このように、第 3 実施形態では、上記第 1 及び第 2 実施形態と同様に、撮影時のプレビュー表示画面において文字認識用文字枠を表示して認識対象の文字画像を撮影し、枠内の画像について文字認識処理を行うことで、文字認識処理にかかる負荷を軽減できる。

また、認識モードの選択によって、予め設定した認識対象文字の種別に応じた文字認識処理を行うことにより、文字認識精度を向上できる。また、認識文字表示画面において、認識結果として複数の種別の文字データを並べて表示し、所定

の文字単位ごとに分けて個別に選択、利用、保存が可能になっている。これにより、例えば電話番号の場合は音声発信、メールアドレスの場合はメール作成など、文字列の種別に応じた認識文字データの利用が容易に可能となる。

5      なお、本発明は上述した実施形態に何ら限定されるものではなく、その要旨を逸脱しない範囲において種々の態様で実施し得るものである。

本発明を詳細にまた特定の実施態様を参照して説明したが、本発明の精神と範囲を逸脱することなく様々な変更や修正を加えることができることは当業者にとって明らかである。

10      本出願は、2002年8月7日出願の日本特許出願No.2002-230312、2003年5月12日出願の日本特許出願No.2003-133171に基づくものであり、その内容はここに参照として取り込まれる。

#### <産業上の利用可能性>

15      以上説明したように本発明によれば、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認識処理方法を提供できる。

## 請 求 の 範 囲

1. 文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、

5 前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、

10 前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手段と、

前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手段と、

を備えた文字認識処理装置。

15 2. 請求の範囲第1項記載の文字認識処理装置であって、

前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを取り込み、

20 前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごとに、前記カーソル位置情報を照合して文字の配置を解析するものである文字認識処理装置。

3. 請求の範囲第1項記載の文字認識処理装置であって、

25 前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを取り込み、

前記レイアウト解析手段は、前記取り込んだ複数の文字画像を連結した画像データについて、前記カーソル位置情報を照合して文字の配置を解析するものである文字認識処理装置。

4. 請求の範囲第1項ないし第3項のいずれか一項に記載の文字認識処理装置を備えた携帯端末装置。

5. 文字認識用の認識対象となる文字画像を撮影する撮影手段と、  
前記撮影された文字画像の画像データを取り込む画像取り込み手段と、  
前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、  
前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、  
前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手段と、  
前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手段と、  
を備えた携帯端末装置。

6. 請求の範囲第5項記載の携帯端末装置であって、  
前記撮影手段による撮影時に、前記文字画像と重ねて文字認識用文字枠を表示する文字枠表示手段を備える携帯端末装置。

7. 請求の範囲第5項記載の携帯端末装置であって、  
前記文字認識手段による認識結果の文字情報を表示する認識文字表示手段を備える携帯端末装置。

8. 請求の範囲第7項記載の携帯端末装置であって、  
前記認識文字表示手段は、前記認識結果の文字情報を所定の文字単位ごとに個別に選択可能に表示するものである携帯端末装置。

9. 請求の範囲第5項記載の携帯端末装置であって、



前記文字認識手段による認識結果の文字情報を保存する認識文字保存手段を備える携帯端末装置。

10. 請求の範囲第9項記載の携帯端末装置であって、

5 前記認識文字保存手段は、前記文字情報を認識文字格納領域に保存するものである携帯端末装置。

11. 請求の範囲第9項記載の携帯端末装置であって、

10 前記認識文字保存手段は、前記文字情報の種別が電話番号、メールアドレス、URL (Uniform Resource Locator) のいずれかであった場合に、その文字情報の種別に応じたデータベースに登録するものである携帯端末装置。

12. 請求の範囲第5項記載の携帯端末装置であって、

15 前記文字認識手段による認識結果の文字情報を当該文字情報の種別に応じて活用する認識文字活用手段を備える携帯端末装置。

13. 請求の範囲第12項記載の携帯端末装置であって、

20 前記認識文字活用手段は、前記文字情報の種別が電話番号であった場合にその電話番号への発信画面を表示するものである携帯端末装置。

14. 請求の範囲第12項記載の携帯端末装置であって、

25 前記認識文字活用手段は、前記文字情報の種別がメールアドレスであった場合にそのメールアドレスへの電子メール作成画面を表示するものである携帯端末装置。

15. 請求の範囲第12項記載の携帯端末装置であって、

前記認識文字活用手段は、前記文字情報の種別がURL (Uniform Resource Locator) であった場合にそのURLへのネットワーク接続画面を表示するものである携帯端末装置。

16. 請求の範囲第5項記載の携帯端末装置であって、  
前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、  
前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数  
5 取得した場合に、これらの複数の文字画像の画像データを連結して取り込み、  
前記文字認識手段は、前記連結された文字画像の画像データの文字認識を行っ  
て文字情報に変換するものである携帯端末装置。

17. 請求の範囲第5項記載の携帯端末装置であって、  
10 前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、  
前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数  
取得した場合に、これらの複数の文字画像の画像データを取り込み、  
前記文字認識手段は、前記取り込まれた複数の文字画像のそれぞれについて画  
像データの文字認識を行って文字情報に変換するものである携帯端末装置。

15

18. 請求の範囲第5項記載の携帯端末装置であって、  
前記文字認識手段は、認識対象の文字種別に応じた複数の認識モードを有し、  
予め設定された認識モードによって対応する文字種別に適した文字認識処理を行  
うものである携帯端末装置。

20

19. 認識対象となる文字画像を文字認識用文字枠とともに表示して撮影  
するステップと、  
前記撮影した文字画像の画像データを取り込むステップと、  
前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情  
25 報を出力するステップと、  
前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合し  
て文字の配置を解析するステップと、  
前記文字配置の解析結果を基に文字画像を切り出すステップと、  
前記切り出された文字画像を文字として認識し文字情報に変換するステップと、

を有することを特徴とする文字認識処理方法。

20. 請求の範囲第19項に記載の各ステップをコンピュータに実行させるための文字認識処理プログラム。

図 1

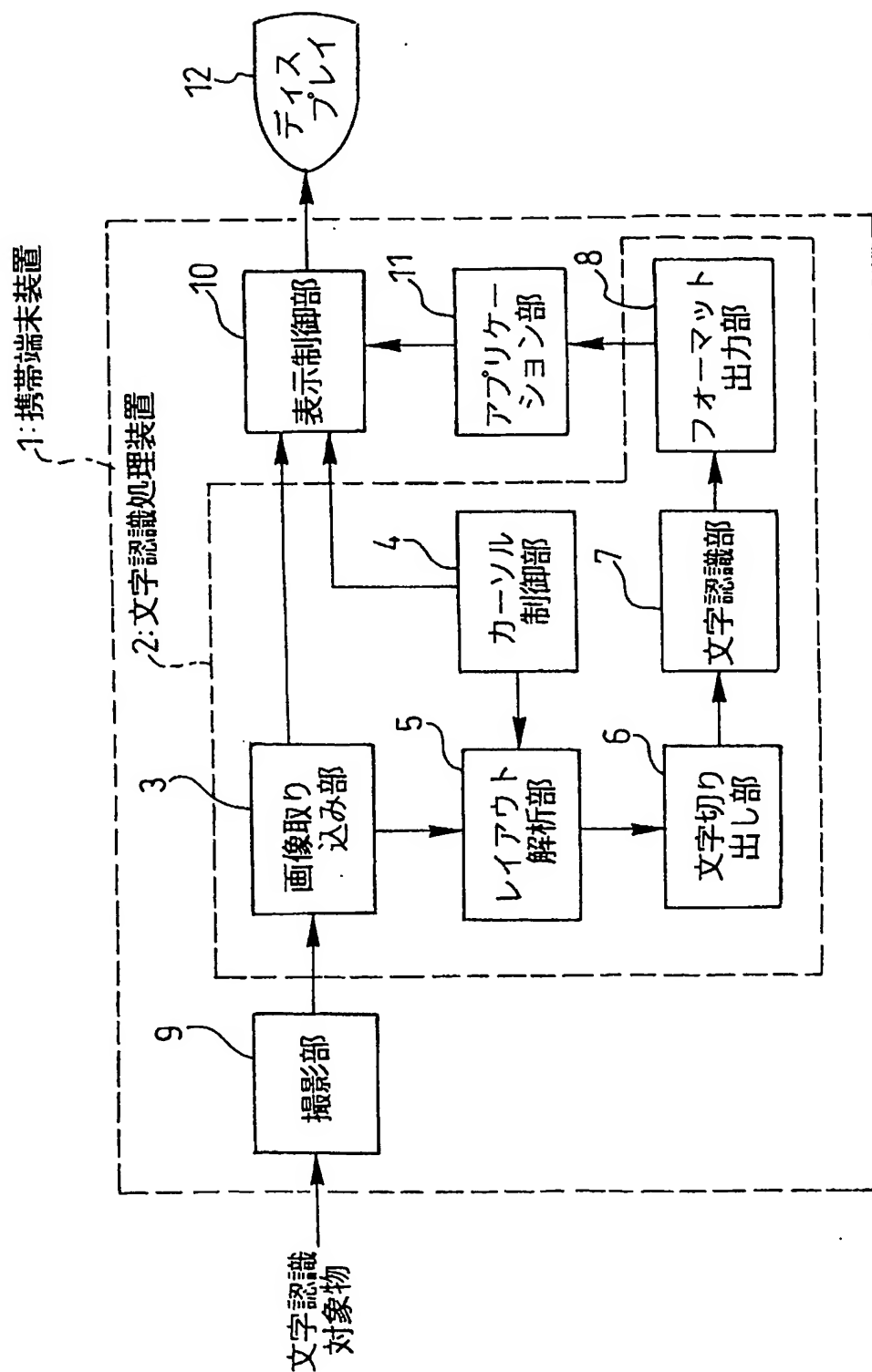


図 2

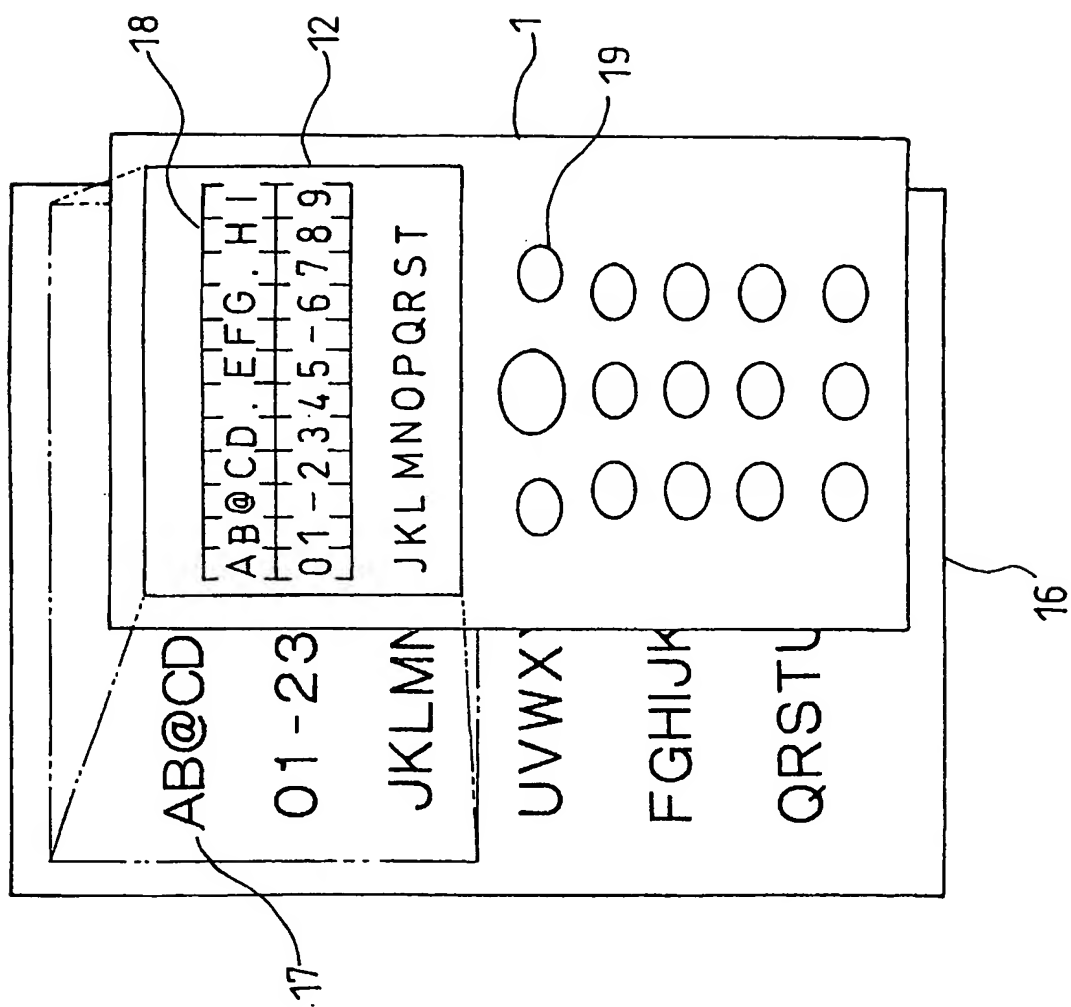


図 3

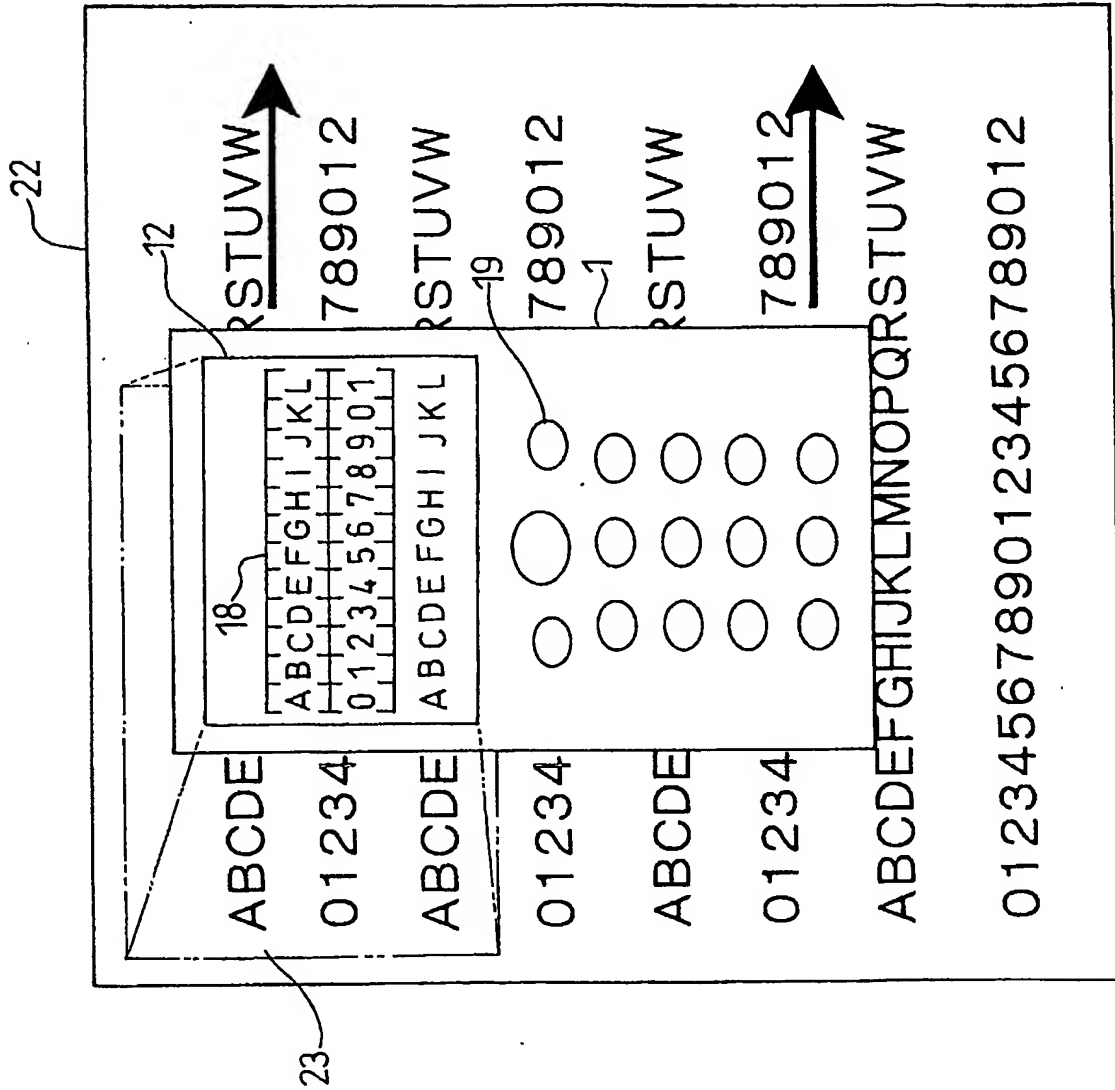


図 4

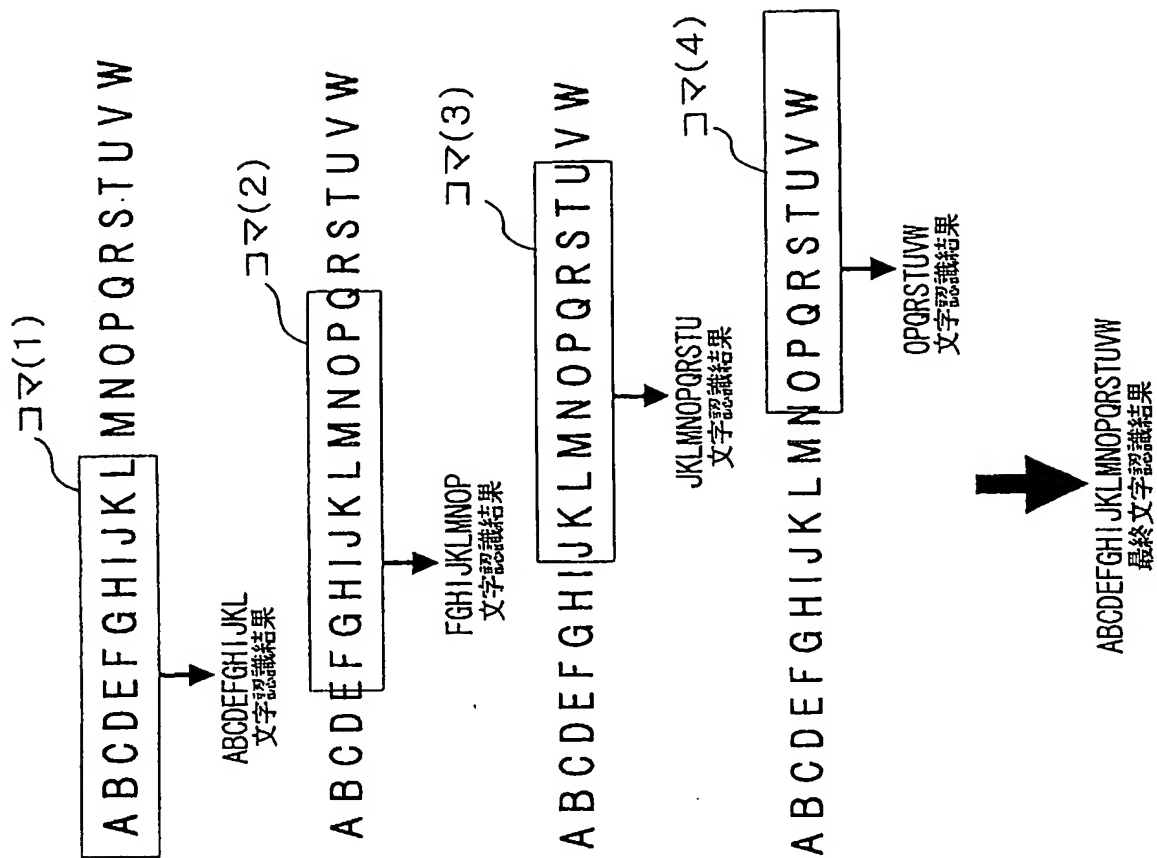


図 5

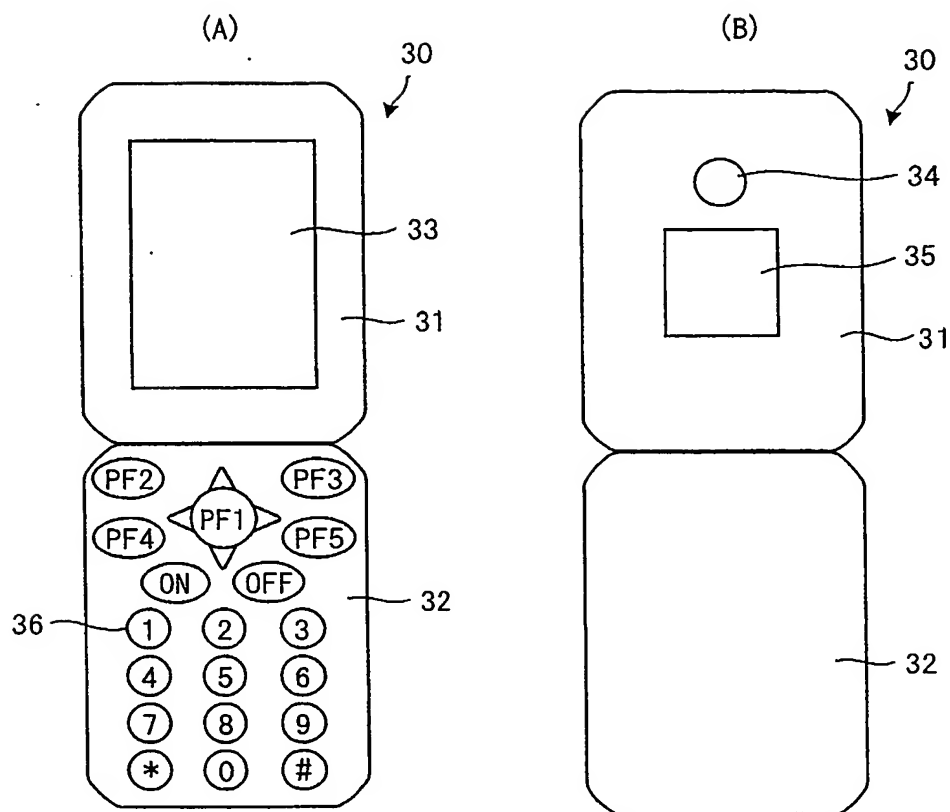




図 6

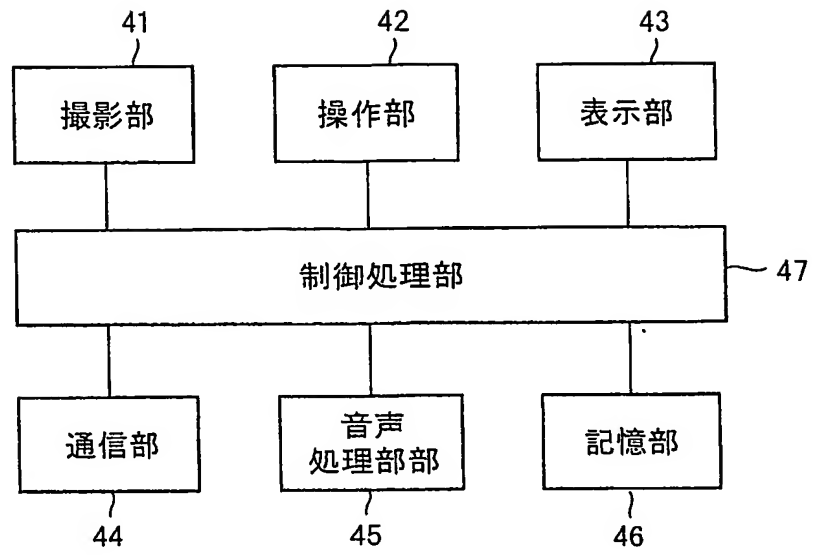


図 7

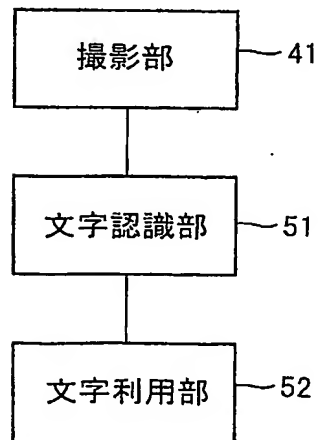


図 8

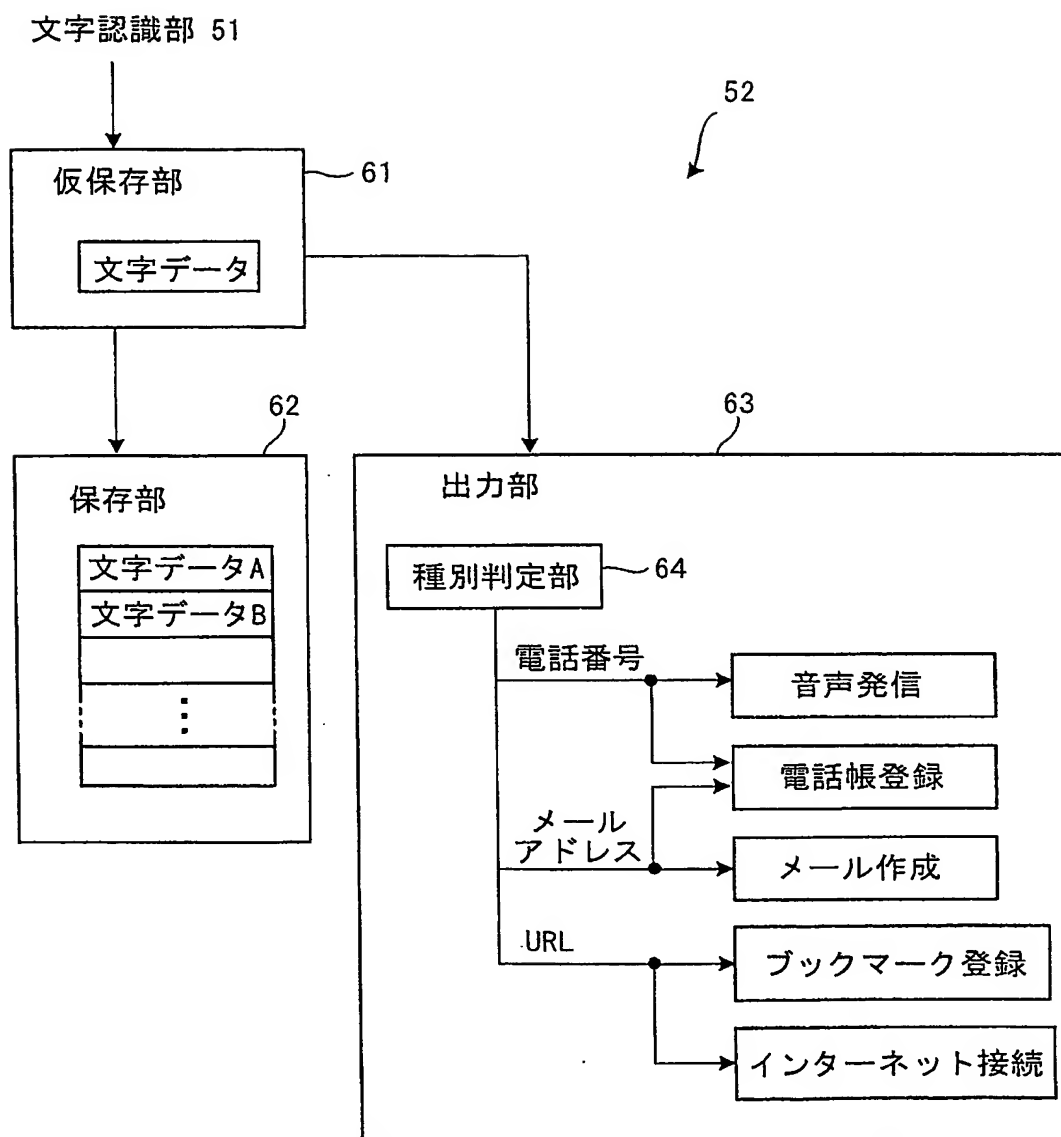


図 9

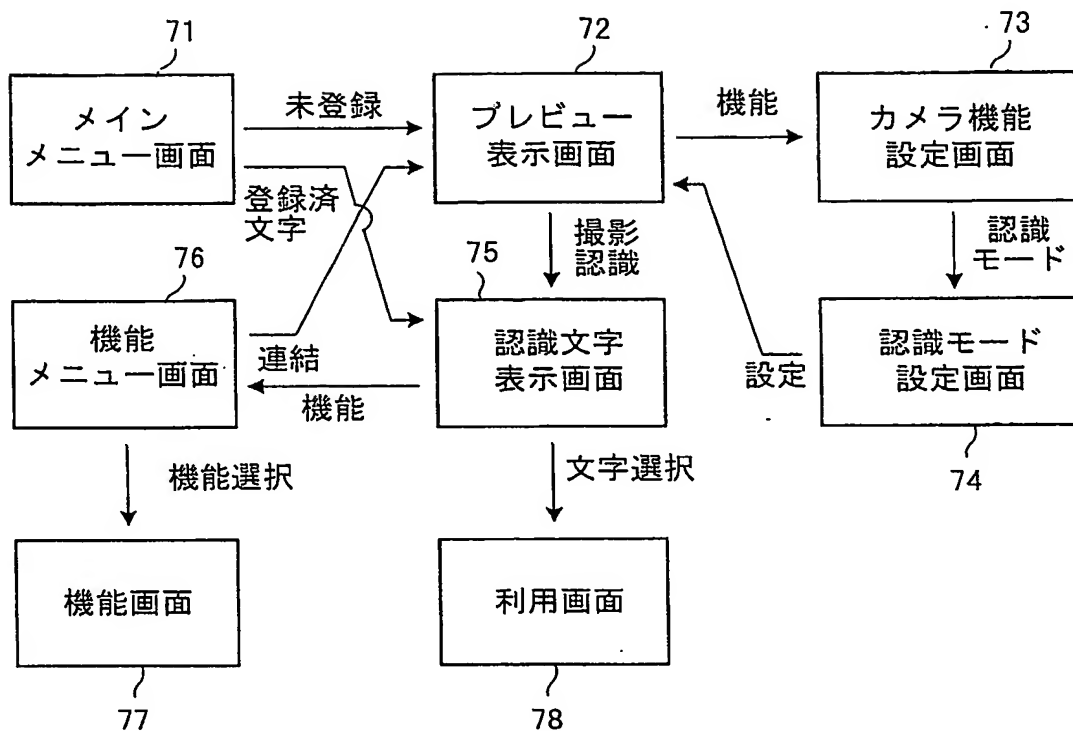


図 10

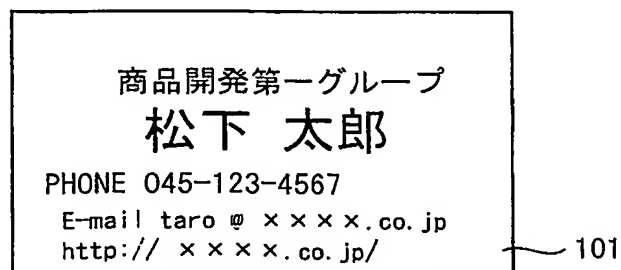


図 1 1

71

メインメニュー
1 未登録
2 未登録
3 09012345678
4 http://〇〇〇〇.com/
5 未登録
6 未登録
7 未登録
8 未登録
選択

図 1 2

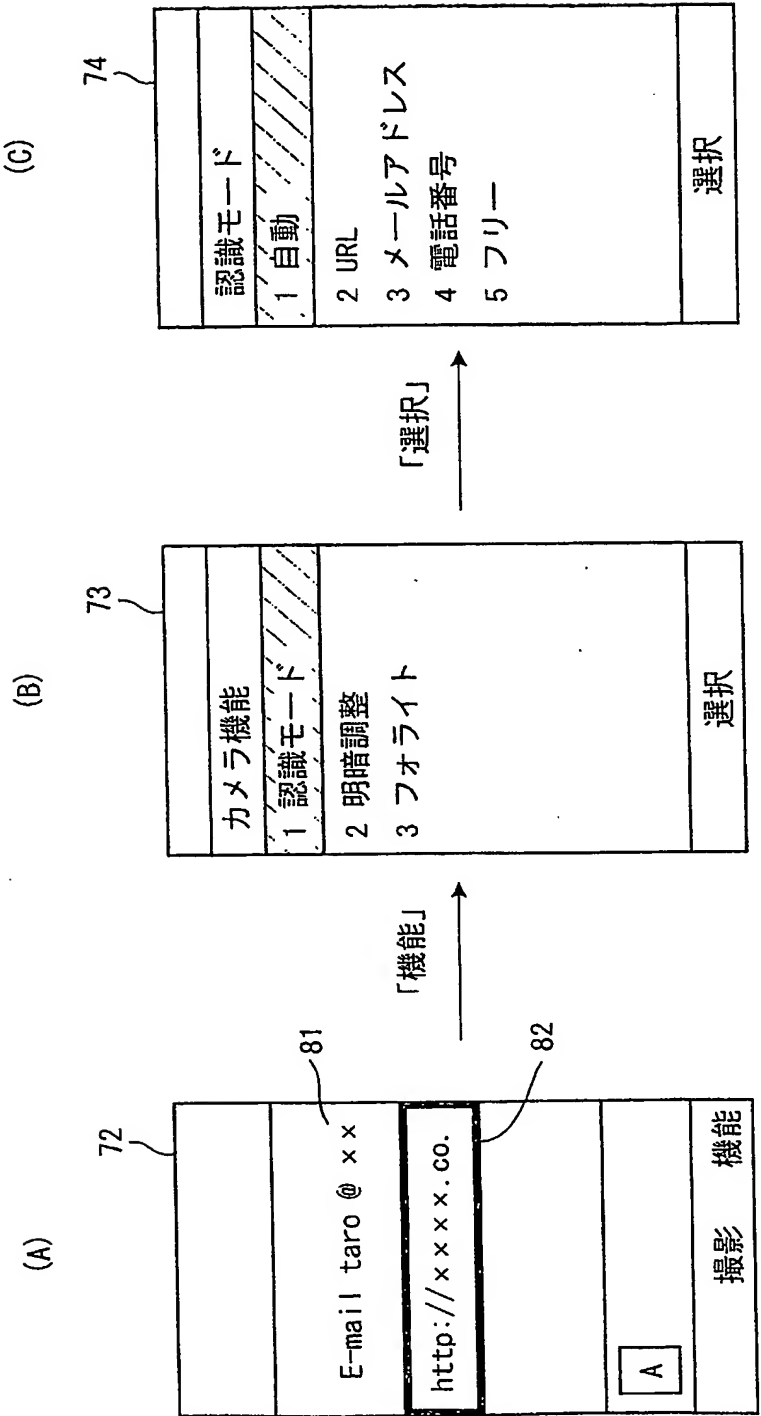
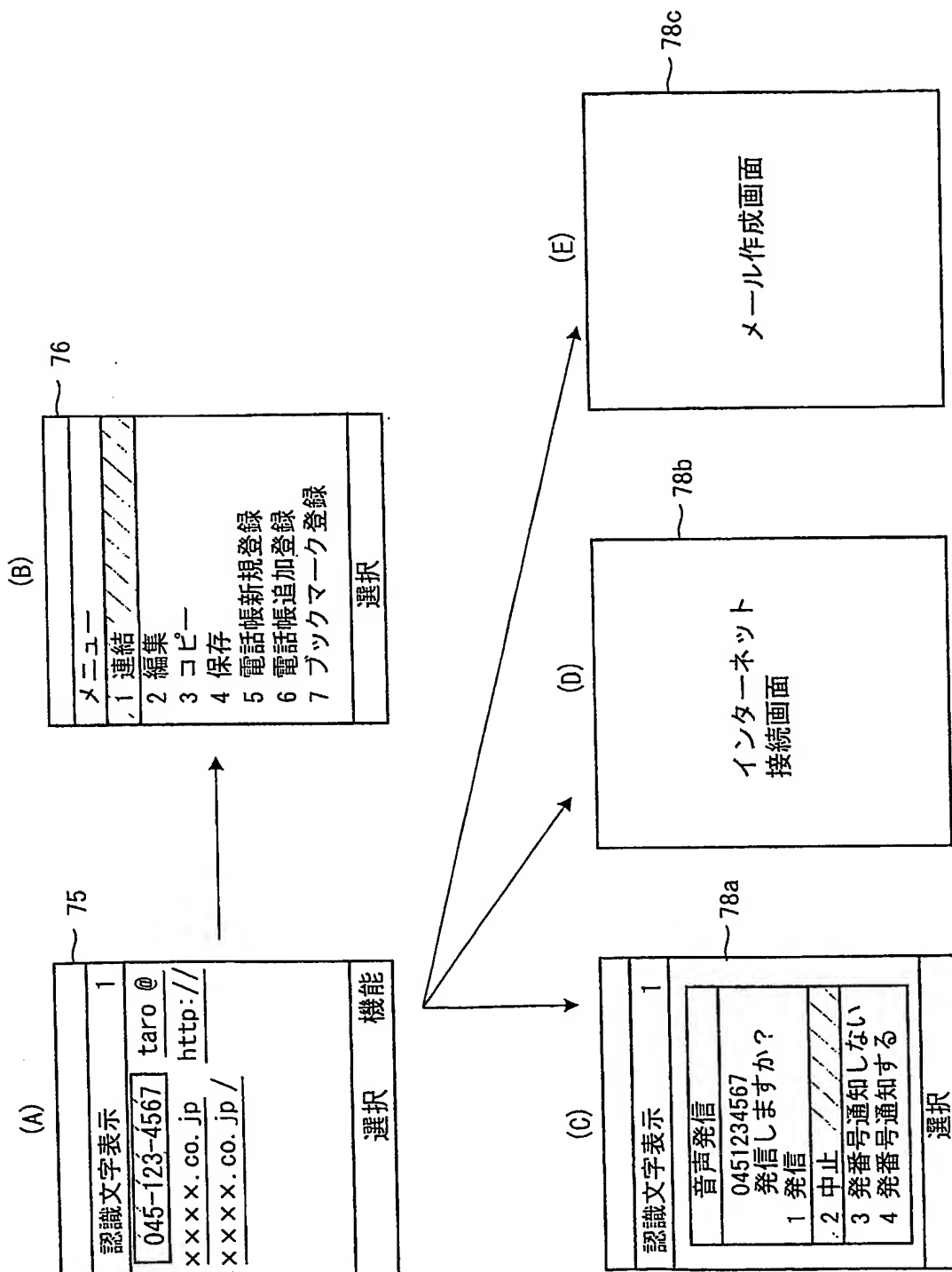


図 1 3



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/09483

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.<sup>7</sup> G06K9/34, H04M1/00, G06K9/03, H04M1/56, G06K9/20

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.<sup>7</sup> G06K9/00-9/82, H04M1/00-56

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Toroku Jitsuyo Shinan Koho	1994-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 9-62786 A (Oki Electric Industry Co., Ltd.), 07 March, 1997 (07.03.97), Full text; Figs. 1 to 3 (Family: none)	1, 19-20 2-18
Y	JP 2002-24762 A (Toshiba Corp.), 25 January, 2002 (25.01.02), Full text; Figs. 1 to 17 & US 2002/12468 A1	2-4, 16-17
Y	JP 7-220025 A (Canon Inc.), 18 August, 1995 (18.08.95), Full text; Figs. 1 to 21 (Family: none)	2-4, 16-17

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>
--	---

Date of the actual completion of the international search  
02 September, 2003 (02.09.03)

Date of mailing of the international search report  
16 September, 2003 (16.09.03)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/09483

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2003-78640 A (Hitachi, Ltd.), 14 March, 2003 (14.03.03), Full text; Figs. 1 to 11 (Family: none)	4-8, 11-12, 14-18
Y	JP 2002-111841 A (Sony Corp.), 12 April, 2002 (12.04.02), Full text; Figs. 1 to 7 (Family: none)	4-6, 11, 13-17
Y	JP 2001-8072 A (Casio Computer Co., Ltd.), 12 January, 2001 (12.01.01), Full text; Figs. 1 to 14 (Family: none)	9-11
A.	JP 2000-330681 A (Sony Corp.), 30 November, 2000 (30.11.00), Full text; Figs. 1 to 12 (Family: none)	11-15



## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G 0 6 K 9 / 3 4 H 0 4 M 1 / 0 0  
 G 0 6 K 9 / 0 3 H 0 4 M 1 / 5 6  
 G 0 6 K 9 / 2 0

## B. 調査を行った分野

## 調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G 0 6 K 9 / 0 0 - 9 / 8 2  
 H 0 4 M 1 / 0 0 - 5 6

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
 日本国公開実用新案公報 1971-2003年  
 日本国実用新案登録公報 1996-2003年  
 日本国登録実用新案公報 1994-2003年

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	J P 9-62786 A (沖電気工業株式会社) 1997. 03. 07, 全文, 図1-3 (ファミリーなし)	1, 19-20 2-18
Y	J P 2002-24762 A (株式会社東芝) 2002. 01. 25, 全文, 図1-17 & US 2002/12468 A1	2-4, 16-17
Y	J P 7-220025 A (キャノン株式会社) 1995. 08. 18, 全文, 図1-21 (ファミリーなし)	2-4, 16-17

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの  
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
 「O」口頭による開示、使用、展示等に言及する文献  
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
 「&」同一パテントファミリー文献

国際調査を完了した日

02. 09. 03

国際調査報告の発送日

16.09.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

廣瀬 文雄

5 H

8526

電話番号 03-3581-1101 内線 3531

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P 2003-78640 A (株式会社日立製作所) 2003. 03. 14, 全文, 図1-11 (ファミリーなし)	4-8, 11-12, 14-18
Y	J P 2002-111841 A (ソニー株式会社) 2002. 04. 12, 全文, 図1-7 (ファミリーなし)	4-6, 11, 13-17
Y	J P 2001-8072 A (カシオ計算機株式会社) 2001. 01. 12, 全文, 図1-14 (ファミリーなし)	9-11
A	J P 2000-330681 A (ソニー株式会社) 2000. 11. 30, 全文, 図1-12 (ファミリーなし)	11-15